

## PEMODELAN BAYESIAN UNTUK ANALISIS DAMPAK PENDIDIKAN TERHADAP TINGKAT PENGANGGURAN GLOBAL

Elvina Benedicta Santoso<sup>1)</sup>, Wilbert Chandra<sup>1)</sup>, Felix Cristiano Bungaran<sup>1)</sup>, Patricia Pepita<sup>1)\*</sup>,  
Ni Luh Putu Satyaning Pradnya Paramita<sup>1)</sup>

<sup>1</sup>Data Science, School of Computer Science, Universitas Bina Nusantara, Indonesia

\*Penulis korespondensi: [patricia.pepita@binus.ac.id](mailto:patricia.pepita@binus.ac.id)

### ABSTRAK

Pengangguran merupakan permasalahan global yang mempengaruhi kondisi ekonomi suatu wilayah atau negara secara signifikan. Studi sebelumnya telah mengidentifikasi faktor-faktor yang mempengaruhi tingkat pengangguran, termasuk aktivitas keuangan, pertumbuhan penduduk, jenis kelamin, status perkawinan, dan pendidikan. Namun, hubungan antara tingkat pendidikan dan pengangguran masih perlu diteliti lebih mendalam karena bisa digunakan untuk memprediksi angka pengangguran. Penelitian ini menggunakan metode statistik Bayesian untuk memahami seberapa jauh tingkat pendidikan mempengaruhi tingkat pengangguran. Data yang berisi 24 variabel independen yang berkaitan dengan tingkat pendidikan dan kepandaian sampel populasi digunakan untuk memprediksi tingkat pengangguran. Model regresi linier beta digunakan dengan algoritma *Markov Chain Monte Carlo* (MCMC) serta metode *Gibbs Sampling*. Riset ini menunjukkan bahwa kemampuan matematika dan membaca di semua tingkat pendidikan secara signifikan mempengaruhi tingkat pengangguran. Kemampuan matematika di tingkat pendidikan dasar (kelas 2 dan 3) serta kemampuan membaca pada akhir pendidikan dasar secara positif mempengaruhi tingkat pengangguran. Sebaliknya, kemampuan membaca di tingkat pendidikan dasar (kelas 2 dan 3), kemampuan matematika pada akhir pendidikan dasar, dan kemampuan matematika pada tingkat pendidikan menengah bawah secara negatif mempengaruhi tingkat pengangguran. Maka dari itu, pendidikan khususnya pendidikan dasar memainkan peran yang signifikan dalam mengurangi tingkat pengangguran. Kemudian, dilihat dari posisi Indonesia di dunia, kemampuan matematika pada pendidikan dasar dan akhir menengah ke bawah di negara Indonesia sedikit kurang dari rata-rata global. Dengan demikian, pendidikan, khususnya kemampuan matematika, perlu ditingkatkan agar dapat menurunkan angka pengangguran di Indonesia. Penelitian ini diharapkan memberikan wawasan yang lebih dalam mengenai pengaruh berbagai tingkat pendidikan terhadap masalah pengangguran.

Kata-kata kunci: Tingkat Pengangguran, Tingkat Pendidikan, Bayesian, Regresi Linier Beta, Signifikan.

### ABSTRACT

*Unemployment is a global problem that significantly affects the economic conditions of a region or country. Previous studies have identified factors that influence unemployment rates, including financial activity, population growth, gender, marital status, and education. However, the link between education level and unemployment requires further study for predictive purposes. This research uses Bayesian statistical methods to understand how far the level of education affects the unemployment rate. Data containing*

24 independent variables related to the level of education of the sample is used to predict the unemployment rate. The beta linear regression model was used with the Markov Chain Monte Carlo (MCMC) algorithm and the Gibbs Sampling method. This research shows that mathematics and reading abilities at all levels of education significantly influence unemployment rates. Mathematics skills at the primary education level (grades 2 and 3) and reading skills at the end of basic education positively influence the unemployment rate. On the other hand, reading ability at the primary education level (grades 2 and 3), mathematics ability at the end of primary education and at the lower secondary education level negatively influence the unemployment rate. Therefore, education, especially primary education, plays a significant role in reducing the unemployment rate. Additionally, mathematics proficiency in primary and lower secondary education in Indonesia falls slightly below the global average. Hence, there is a need to enhance education, particularly in mathematics, to alleviate unemployment in Indonesia. This study aims to provide insights into how education levels affect unemployment rates.

*Keyword:* Unemployment Rate, Education Level, Bayesian, Beta Linear Regression, Significant.

## Pendahuluan

Pengangguran merupakan suatu permasalahan global karena memberikan pengaruh yang signifikan terhadap kondisi ekonomi suatu wilayah atau negara (Chakraborty et al. 2020). Menurut Sukirno yang dikutip oleh Bachtiar (2019), pengangguran sendiri dapat diartikan sebagai keadaan dimana seseorang tidak memiliki pekerjaan meskipun masih dalam usia produktif. Tingkat pengangguran di suatu negara atau wilayah dipengaruhi oleh berbagai faktor yang bervariasi sesuai kondisi wilayahnya.

Penelitian sebelumnya telah menganalisis faktor-faktor yang menyebabkan tingginya angka pengangguran. Penelitian pada tahun 2020 yang dilakukan oleh Shabbir et al. (2020) menemukan bahwa terdapat relasi positif antara aktivitas keuangan (kredit) dan pertumbuhan penduduk terhadap tingkat pengangguran di Asia Selatan. Selanjutnya, pada tahun yang sama, Aprilliofany (2020) melihat bahwa faktor jenis kelamin, status dalam rumah tangga, status perkawinan, klasifikasi wilayah, dan klasifikasi jurusan akademik mempengaruhi tingkat pengangguran di wilayah Jawa Barat dan Garut. Kemudian, Prakoso (2020) menemukan bahwa faktor pendidikan, inflasi, UMK, dan rata-rata lama sekolah memiliki pengaruh yang signifikan dan negatif terhadap masalah pengangguran di rentang tahun 2010 hingga 2018. Artinya, semakin tinggi nilai variabel tersebut, tingkat pengangguran semakin berkurang.

Meski telah terdapat penelitian mengenai hubungan tingkat pendidikan terhadap pengangguran, belum terdapat penelitian yang lebih mendetail terkait hubungan keduanya. Padahal, hal ini penting untuk diteliti karena bisa digunakan untuk memprediksi angka pengangguran di masa mendatang. Pendidikan menurut KBBI adalah sebuah metode, cara, ataupun tindakan membimbing. Tingkat pendidikan pada umumnya dimulai dari tingkat dasar, menengah, hingga perguruan tinggi (Pristiwanti et al. 2022). Oleh karena itu, diperlukan penelitian lebih lanjut mengenai pengaruh level edukasi terhadap tingkat pengangguran.

Dalam riset ini, peneliti akan menggunakan teknik analisis dengan metode statistik Bayesian untuk mencari seberapa jauh tingkat pendidikan mempengaruhi tingkat pengangguran. Statistik Bayesian merupakan pendekatan statistik yang dianggap sebagai ukuran keyakinan maupun ketidakpastian terhadap suatu kejadian (Populix 2024). Terdapat beberapa konsep statistik Bayesian menurut artikel Populix (2024), diantaranya adalah probabilitas awal, peluang, probabilitas posterior, teorema Bayes, dan inferensi Bayesian. Dalam menggunakan statistik Bayesian, kita perlu mengetahui jenis distribusi dalam pengambilan suatu data. Secara umum, terdapat dua tipe data, yaitu diskrit dan kontinu (Zangre 2023). Dikutip dari sumber yang sama, data merupakan nilai yang tetap dan dapat dihitung sedangkan kontinu

memiliki nilai yang berkelanjutan. Setiap data memiliki distribusi yang berbeda tergantung dari tipe data dan karakteristik dari data tersebut. Beberapa distribusi data meliputi distribusi normal, binomial, poisson, beta, dan gamma (Hussain 2023).

Peneliti akan menggunakan data set yang memiliki 24 variabel bebas yang berkaitan dengan tingkat pendidikan dan kepandaian sampel populasi. Variabel targetnya adalah tingkat pengangguran yang memiliki nilai antara 0 dengan 1. Oleh karena itu, jenis distribusi yang akan dipakai untuk mencari prediksi dari hasil tingkat pengangguran adalah beta regresi. Algoritma yang akan dipakai yaitu *Markov Chain Monte Carlo* (MCMC), di mana pengambilan sampel akan berdasarkan metode *Gibbs Sampling*. *Gibbs Sampling* merupakan sebuah cara pengambilan sampel dari dua atau lebih dimensi atau distribusi multivariat (Zvornicanin 2023). Dalam penelitian ini, sampel awal akan dibuang untuk menghindari ketidaktelitian dalam merepresentasikan distribusi yang diinginkan. Peneliti ingin agar sampel yang dipakai menjadi bebas atau tidak bergantung satu sama lain sehingga hasil dapat menjadi konvergen atau mewakili distribusi data dengan tepat.

Riset ini bertujuan untuk mengidentifikasi dan memahami pengaruh tingkat edukasi terhadap tingkat pengangguran secara mendalam. Dengan menggunakan berbagai variabel mengenai tingkat pendidikan dan kepandaian yang mendetail, kajian ini diharapkan dapat melengkapi dan memperdalam pengetahuan mengenai tingkat pendidikan sebagai faktor penyebab masalah pengangguran.

## Metode

Kajian ini dilakukan di Universitas Bina Nusantara pada tanggal 15 Februari 2024. Bahan yang digunakan pada kajian ini adalah kumpulan data yang diambil dari situs web *kaggle* berjudul *World Educational Data* melalui link berikut [World Educational Data \(kaggle.com\)](https://www.kaggle.com/datasets/wesamzayed/world-educational-data). Dataset yang digunakan memiliki 202 baris dan 29 kolom. Variabel yang terdapat pada dataset adalah Latitude, Longitude, OOSR\_Pre0Primary\_Age\_Male, OOSR\_Pre0Primary\_Age\_Female, OOSR\_Primary\_Age\_Male, OOSR\_Primary\_Age\_Female, OOSR\_Lower\_Secondary\_Age\_Male, OOSR\_Lower\_Secondary\_Age\_Female, OOSR\_Upper\_Secondary\_Age\_Male, OOSR\_Upper\_Secondary\_Age\_Female, Completion\_Rate\_Primary\_Male, Completion\_Rate\_Primary\_Female, Completion\_Rate\_Lower\_Secondary\_Male, Completion\_Rate\_Lower\_Secondary\_Female, Completion\_Rate\_Upper\_Secondary\_Male, Completion\_Rate\_Upper\_Secondary\_Female, Grade\_2\_3\_Proficiency\_Reading, Grade\_2\_3\_Proficiency\_Math, Primary\_End\_Proficiency\_Reading,

Primary\_End\_Proficiency\_Math, Lower\_Secondary\_End\_Proficiency\_Reading, Lower\_Secondary\_End\_Proficiency\_Math, Youth\_15\_24\_Literacy\_Rate\_Male, Youth\_15\_24\_Literacy\_Rate\_Female, Birth\_Rate, Gross\_Primary\_Education\_Enrollment, Gross\_Tertiary\_Education\_Enrollment, Unemployment\_Rate. Variabel Unemployment\_Rate akan digunakan sebagai variabel target dan 24 variabel yang berkaitan dengan tingkat pendidikan digunakan sebagai *independent features* atau variabel bebas. Alat yang digunakan pada kajian ini adalah *software R Studio* dengan *library RJAGS*.

Tahapan kajian :

- Memilih *dataset* yang sesuai dengan judul dan masalah yang ingin diselesaikan
- Pengambilan data-data yang bersifat numerik
- Melakukan *data exploration* untuk mendapatkan statistik deskriptif dan deskripsi sederhana data
- Memilih *features* yang akan digunakan
- Melakukan *feature engineering* dengan melakukan penskalaan pada variabel target dan mengisi nilai kosong pada data berdasarkan median data karena distribusi data yang tidak normal
- Membuat model *Beta Linear Regression* untuk mendapatkan efek dari setiap *features* terhadap variabel target
- Menganalisis *summary* model dan mengambil kesimpulan berdasarkan tingkat kepercayaan 95%
- Membuat dan menganalisa *plot* korelasi antar variabel yang signifikan dengan variabel target secara global dan di negara Indonesia dengan tidak mengikutsertakan nilai *null* pada data.

Dalam rangka menemukan faktor-faktor edukasi yang berdampak pada tingkat pengangguran di suatu negara, pada analisis ini digunakan model *Beta Linear Regression*. *Beta[1]* merupakan *intercept* yang mana merupakan nilai dari variabel dependen saat semua variabel independen bernilai 0. Kemudian, *Beta[2, n]* merupakan koefisien dari setiap *features* dalam model regresi. Sehingga, *Beta[2, n]* menggambarkan seberapa besar dampak relatif dari masing-masing fitur terhadap variabel dependen dalam model regresi.

Variabel respon dalam analisis ini merupakan tingkat pengangguran atau *unemployment rate* (UNP) yang mana nilainya berada dalam rentang 0 hingga 1. Misalkan  $UNP_{ij}$  adalah persentase pengangguran di suatu negara j pada sampel i, berikut rumus matematisnya:

$$Y_i = \frac{UNP_{ij}}{100} \in [0,1]$$

Model regresi yang digunakan pada analisis ini adalah sebagai berikut:

$$\text{logit}(q_i) = \sum_{j=1}^p X_{ij} B_j$$

Dengan  $Likelihood Y_i \sim Beta(rq_i, r(1 - q_i))$  agar nilai  $Y_i$  adalah  $q_i \in [0,1]$  dan konsentrasi sekitar  $q_i$  ditentukan oleh  $r > 0$ . Di mana  $X_{ij}$  adalah nilai dari setiap *features* untuk setiap iterasi  $i$  dengan koefisien regresi mempunyai *noninformative priors*  $B_j \sim Normal(0, 10^2)$  dan parameter konsentrasi mempunyai *prior*  $r \sim Gamma(0.1, 0.1)$ . Dengan menggunakan model regresi tersebut, *Stochastic Search Variable Selection (SSVS)* digunakan untuk membantu mengidentifikasi fitur-fitur (variabel) mana yang sangat berdampak pada tingkat pengangguran.

Model dibuat dengan menggunakan *package JAGS* pada bahasa pemrograman R. Untuk membangun model, 10000 sampel pertama akan diabaikan (*burn-in*) dan 75000 sampel berikutnya diambil dari masing-masing 2 *chains* dengan *thinning factor* sebesar 5. Kekonvergenan model dinilai dengan memeriksa *trace-plot* dan mempertimbangkan *output* dari *Gelman-Rubin statistic* dan *effective sample size* (EES).

*Trace plot* dapat menunjukkan suatu model sudah konvergen atau belum. Model dikatakan sudah konvergen apabila *trace plot* menunjukkan nilai-nilai variabel acak stabil. Selain itu, *plot* yang rapat menunjukkan bahwa hasil sampling stabil. Berdasarkan *Gelman-Rubin statistic*, suatu model dikatakan konvergen apabila nilai yang didapatkan kurang dari 1.1, sedangkan dengan menggunakan EES suatu model diindikasikan konvergen apabila nilai yang didapatkan lebih dari 1000.

*Scatterplot* dan *bar chart* dibuat menggunakan *library plotly* pada bahasa pemrograman Python. *Scatterplot* merepresentasikan variabel yang signifikan dan variabel target. Sumbu x merepresentasikan variabel yang berperan penting dan sumbu y adalah tingkat pengangguran terhadap variabel tersebut. Kemudian, *bar chart* merepresentasikan perbandingan antara negara Indonesia dan negara di dunia pada variabel yang signifikan dan variabel target. Sumbu x merupakan variabel yang dianalisis dan sumbu y merupakan tingkat atau nilai dari variabel tersebut.

## Hasil dan Pembahasan

Pada model yang telah dibuat, *trace plot* memperlihatkan konvergen pada model. Komputasi *Gelman-Rubin statistic* pada setiap parameter dalam model mendapatkan *multivariate potential scale reduction factors* (PSRF) 1.01. *Effective sample size* (ESS) untuk setiap parameter yang telah dikomputasikan melebih 1000. Dengan ini model yang dibuat telah mencapai konvergensi.

Posterior dari parameter dirangkum dalam Tabel 1. Berdasarkan hasil dari SSVS, seluruh semua parameter signifikan sehingga dapat disimpulkan bahwa dampak edukasi terhadap tingkat pengangguran jika setiap *feature* dianggap tetap adalah sebagai berikut:

- Setiap pertambahan 1 unit pada *Out-of-school rate for pre-primary age males* mengalikan probabilitas pengangguran sebanyak  $\exp(-0.039)$ .
- Setiap pertambahan 1 unit pada *Out-of-school rate for pre-primary age females* mengalikan probabilitas pengangguran sebanyak  $\exp(0.16)$ .
- Setiap pertambahan 1 unit pada *Out-of-school rate for primary age males* mengalikan probabilitas pengangguran sebanyak  $\exp(0.073)$ .
- Setiap pertambahan 1 unit pada *Out-of-school rate for primary age females* mengalikan probabilitas pengangguran sebanyak  $\exp(0.002)$ .
- Setiap pertambahan 1 unit pada *Out-of-school rate for lower secondary age males* mengalikan probabilitas pengangguran sebanyak  $\exp(0.289)$ .
- Setiap pertambahan 1 unit pada *Out-of-school rate for lower secondary age females* mengalikan probabilitas pengangguran sebanyak  $\exp(-0.314)$ .
- Setiap pertambahan 1 unit pada *Out-of-school rate for upper secondary age males* mengalikan probabilitas pengangguran sebanyak  $\exp(-0.328)$ .
- Setiap pertambahan 1 unit pada *Out-of-school rate for upper secondary age females* mengalikan probabilitas pengangguran sebanyak  $\exp(0.159)$ .
- Setiap pertambahan 1 unit pada *Completion rate for primary education among males* mengalikan probabilitas pengangguran sebanyak  $\exp(0.256)$ .
- Setiap pertambahan 1 unit pada *Completion rate for primary education among females* mengalikan probabilitas pengangguran sebanyak  $\exp(-0.264)$ .
- Setiap pertambahan 1 unit pada *Completion rate for lower secondary education among males* mengalikan probabilitas pengangguran sebanyak  $\exp(-0.191)$ .
- Setiap pertambahan 1 unit pada *Completion rate for lower secondary education among females* mengalikan probabilitas pengangguran sebanyak  $\exp(0.362)$ .
- Setiap pertambahan 1 unit pada *Completion rate for upper secondary education among males* mengalikan probabilitas pengangguran sebanyak  $\exp(-0.014)$ .
- Setiap pertambahan 1 unit pada *Completion rate for upper secondary education among females* mengalikan probabilitas pengangguran sebanyak  $\exp(0.002)$ .
- Setiap pertambahan 1 unit *Proficiency in reading for grade 2-3 students* mengalikan probabilitas pengangguran sebanyak  $\exp(-0.113)$ .

- Setiap pertambahan 1 unit *Proficiency in math for grade 2-3 students* mengalikan probabilitas pengangguran sebanyak  $\exp(0.146)$ .
- Setiap pertambahan 1 unit *Proficiency in reading at the end of primary education* mengalikan probabilitas pengangguran sebanyak  $\exp(0.085)$ .
- Setiap pertambahan 1 unit *Proficiency in math at the end of primary education* mengalikan probabilitas pengangguran sebanyak  $\exp(-0.123)$ .
- Setiap pertambahan 1 unit *Proficiency in reading at the end of lower secondary education* mengalikan probabilitas pengangguran sebanyak  $\exp(0.102)$ .
- Setiap pertambahan 1 unit *Proficiency in math at the end of lower secondary education* mengalikan probabilitas pengangguran sebanyak  $\exp(-0.143)$ .
- Setiap pertambahan 1 unit *Literacy rate among male youths aged 15-24* mengalikan probabilitas pengangguran sebanyak  $\exp(-0.094)$ .
- Setiap pertambahan 1 unit *Literacy rate among female youths aged 15-24* mengalikan probabilitas pengangguran sebanyak  $\exp(0.086)$ .
- Setiap pertambahan 1 unit *Gross enrollment in primary education* probabilitas pengangguran sebanyak  $\exp(-0.008)$ .
- Setiap pertambahan 1 unit *Gross enrollment in tertiary education* probabilitas pengangguran sebanyak  $\exp(0.027)$ .

Tabel 1. Rangkuman nilai posterior

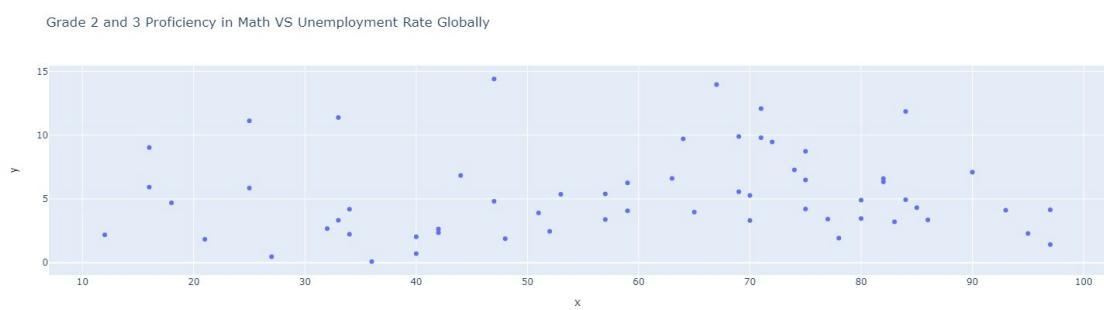
Parameter	Variabel yang Mewakili	Estimasi	95% Tingkat Kepercayaan
$\beta_0$	titik potong/ <i>intercept</i>	-2.684	[-2.784, -2.585]
$\beta_1$	OOSR_Pre0Primary_Age_Male	-0.039	[-0.507, 0.447]
$\beta_2$	OOSR_Pre0Primary_Age_Female	0.16	[-0.312, 0.615]
$\beta_3$	OOSR_Primary_Age_Male	0.073	[-0.332, 0.477]
$\beta_4$	OOSR_Primary_Age_Female	0.002	[-0.455, 0.461]
$\beta_5$	OOSR_Lower_Secondary_Age_Male	0.289	[-0.097, 0.662]
$\beta_6$	OOSR_Lower_Secondary_Age_Female	-0.314	[-0.733, 0.113]

$\beta_7$	OOSR_Upper_Secondary_Age_Male	-0.328	[-0.753, 0.079]
$\beta_8$	OOSR_Upper_Secondary_Age_Female	0.159	[-0.281, 0.598]
$\beta_9$	Completion_Rate_Primary_ Male	0.256	[-0.304, 0.788]
$\beta_{10}$	Completion_Rate_Primary_ Female	-0.264	[-0.789, 0.283]
$\beta_{11}$	Completion_Rate_Lower_ Secondary_Male	-0.191	[-0.812, 0.413]
$\beta_{12}$	Completion_Rate_Lower_ Secondary_Female	0.362	[-0.325, 1.046]
$\beta_{13}$	Completion_Rate_Upper_ Secondary_Male	-0.014	[-0.474, 0.444]
$\beta_{14}$	Completion_Rate_Upper_ Secondary_Female	0.002	[-0.514, 0.522]
$\beta_{15}$	Grade_2_3_Proficiency_ Reading	-0.113	[-0.245, 0.024]
$\beta_{16}$	Grade_2_3_Proficiency_Math	0.146	[0.027, 0.264]
$\beta_{17}$	Primary_End_Proficiency_ Reading	0.085	[-0.07, 0.231]
$\beta_{18}$	Primary_End_Proficiency_ Math	-0.123	[-0.299, 0.058]
$\beta_{19}$	Lower_Secondary_End_ Proficiency_Reading	0.102	[-0.087, 0.292]
$\beta_{20}$	Lower_Secondary_End_ Proficiency_Math	-0.143	[-0.337, 0.06]
$\beta_{21}$	Youth_15_24_Literacy_Rate_Male	-0.094	[-0.679, 0.497]
$\beta_{22}$	Youth_15_24_Literacy_Rate_ Female	0.086	[-0.511, 0.704]
$\beta_{23}$	Gross_Primary_Educational_ Enrollment	-0.008	[-0.138, 0.135]

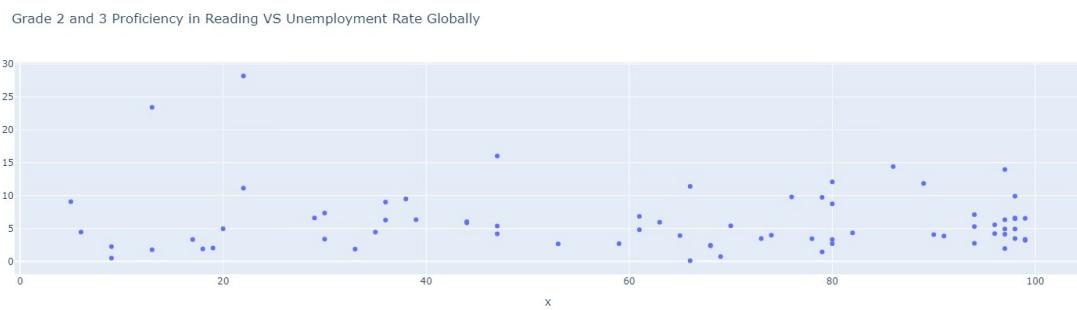
$\beta_{24}$	Gross_Tertiary_Educational_ Enrollment	0.027	[ -0.116, 0.169 ]
--------------	--	-------	----------------------

Dari tabel di atas, variabel-variabel yang signifikan terhadap tingkat pengangguran *grade 2-3 proficiency in math*, *grade 2-3 proficiency in reading*, *primary end proficiency in math*, *primary end proficiency in reading*, dan *lower secondary end proficiency in math*. Untuk melihat persebaran data dengan lebih baik dan keperluan visualisasi, variabel tersebut akan di-plot menjadi berikut ini:

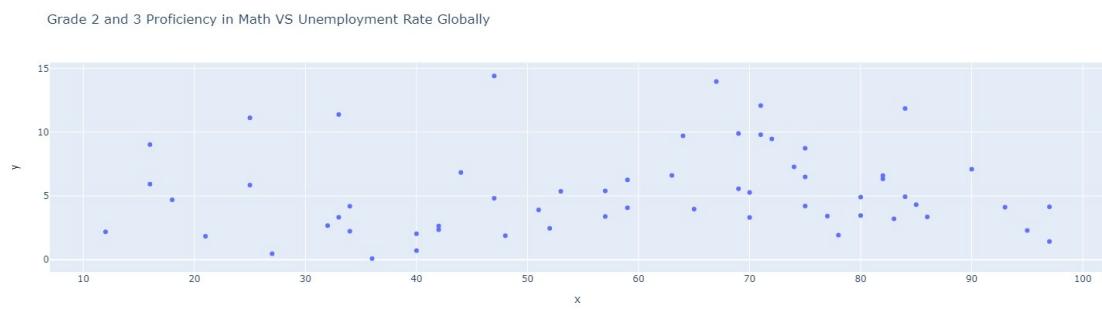
- Persebaran variabel yang signifikan terhadap variabel target.



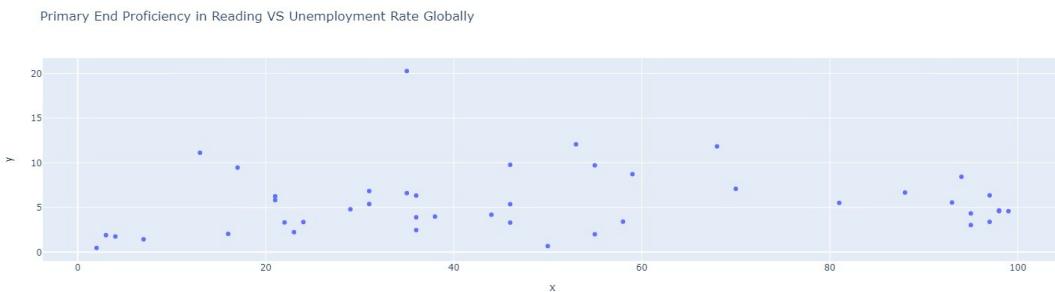
Gambar 1. Scatterplot dari variabel *grade 2-3 proficiency in math* terhadap *unemployment rate*



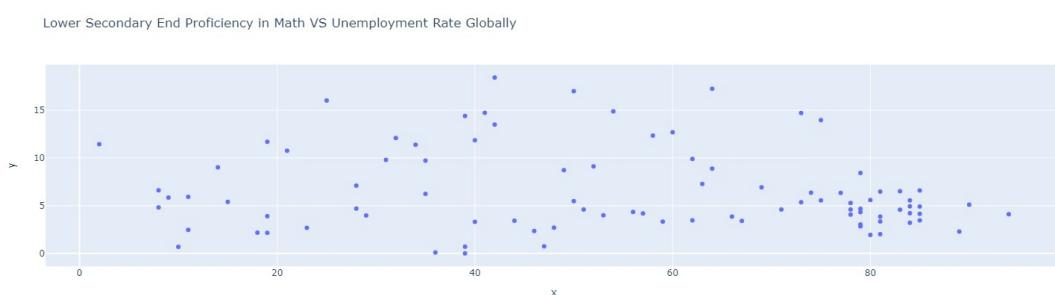
Gambar 2. Scatterplot dari variabel *grade 2-3 proficiency in reading* terhadap *unemployment rate*



Gambar 3. Scatterplot dari variabel *primary end proficiency in math* terhadap *unemployment rate*



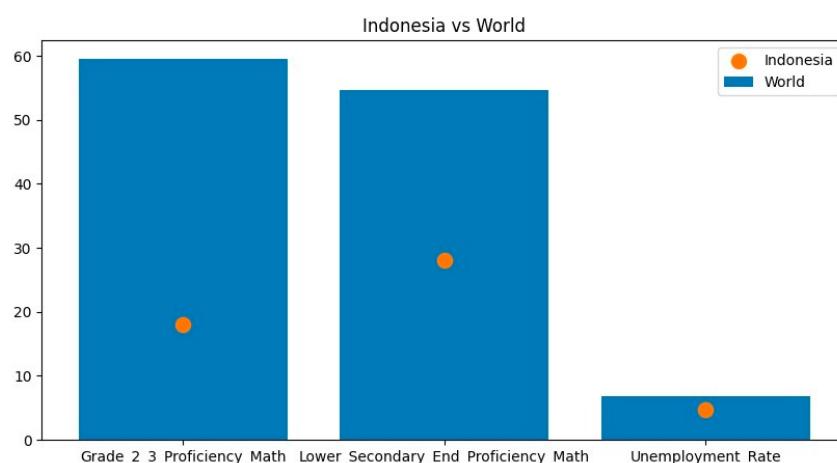
Gambar 4. Scatterplot dari variabel *primary end proficiency in reading* terhadap *unemployment rate*



Gambar 5. Scatterplot dari variabel *lower secondary end proficiency in math* terhadap *unemployment rate*

*Scatterplot* di atas dibuat dengan tidak mengikutsertakan nilai *null* value. Secara keseluruhan, *plot* menunjukkan bahwa semakin tinggi nilai *proficiency*, semakin rendah titik tingkat penganggurannya.

- Perbandingan antara negara Indonesia dan negara di dunia pada dua variabel yang signifikan dan variabel target.



Gambar 6. Bar chart dan titik data dari perbandingan dua variabel yang signifikan terhadap *unemployment rate* di Indonesia dan di dunia.

*Bar chart* dibuat dengan tidak mengikutsertakan data *null*. Tiga variabel yang signifikan lainnya, yaitu *grade 2-3 proficiency reading*, *primary end proficiency reading*, dan *primary end proficiency math* tidak di *plotting* karena data negara Indonesia *null* sehingga tidak dapat merepresentasikan posisi Indonesia di dunia dengan baik. Gambar di atas merepresentasikan bahwa tingkat pengangguran di Indonesia mendekati batas atas rata-rata tingkat pengangguran global. Hal ini didukung oleh titik data Indonesia yang berada jauh di bawah batas atas kedua variabel yang signifikan.

### Kesimpulan

Berdasarkan *slope* dan tingkat kepercayaan 95% dari semua fitur yang digunakan, *summary* model menunjukkan bahwa variabel pendidikan yang signifikan terhadap tingkat pengangguran adalah *grade 2-3 proficiency in math*, *grade 2-3 proficiency in reading*, *primary end proficiency in math*, *primary end proficiency in reading*, dan *lower secondary end proficiency in math*. Hal ini menunjukan bahwa kemampuan matematika dan membaca di semua tingkat pendidikan secara signifikan mempengaruhi tingkat pengangguran. Kemampuan matematika di tingkat pendidikan dasar (kelas 2 dan 3) serta kemampuan membaca pada akhir pendidikan dasar secara positif mempengaruhi tingkat pengangguran. Sebaliknya, kemampuan membaca di tingkat pendidikan dasar (kelas 2 dan 3), kemampuan matematika pada akhir pendidikan dasar, dan kemampuan matematika pada tingkat pendidikan menengah bawah secara negatif mempengaruhi tingkat pengangguran.

Sehingga dapat disimpulkan bahwa faktor yang paling umum berpengaruh pada angka pengangguran di suatu negara adalah kemampuan membaca dan matematika pendidikan dasar di negara tersebut. Maka dari itu, pendidikan khususnya pendidikan dasar memainkan peran yang sangat penting dalam mengurangi tingkat pengangguran. Meningkatkan kualitas pendidikan dasar akan membantu menekan tingkat pengangguran secara global. Hal ini juga didukung dengan keseluruhan persebaran data pada *scatterplot* yang dibuat, dimana semakin tinggi nilai *proficiency*, semakin rendah tingkat penganggurnya.

Terakhir, melalui *barchart* yang dibuat, peneliti menemukan bahwa tingkat pengangguran di Indonesia cukup tinggi karena melebihi rata-rata tingkat pengangguran global. Di sisi lain, titik data merepresentasikan bahwa kemampuan matematika pada pendidikan dasar dan akhir menengah ke bawah di negara Indonesia sedikit kurang dari rata-rata global. Dengan demikian, pendidikan, khususnya kemampuan matematika, perlu ditingkatkan agar dapat menurunkan angka pengangguran di Indonesia.

### **Ucapan Terima Kasih**

Puji dan syukur kami haturkan kepada Tuhan Yang Maha Esa, karena berkat, pengajaran, dan kasih-Nya yang besar kepada para penulis sehingga bisa menyelesaikan penulisan artikel ilmiah PKM-AI yang berjudul “PEMODELAN BAYESIAN UNTUK ANALISIS DAMPAK PENDIDIKAN TERHADAP TINGKAT PENGANGGURAN GLOBAL” tepat pada waktunya.

Penulis ingin mengucapkan terima kasih kepada Ibu Ni Luh Putu Satyaning Pradnya Paramita, S.Si., M.Si., Ph.D selaku dosen pembimbing yang bersedia meluangkan waktu, tenaga, serta pikiran, dan turut memberi perhatian dalam penulisan artikel ilmiah ini.

Terima kasih kami sampaikan juga kepada pihak BINUS University atas kesempatan yang diberikan sehingga kami bisa melewati tahap seleksi dalam kampus. Tidak lupa juga kami mengucapkan terima kasih kepada pihak SIMBELMAWA atas kesempatan yang diberikan sehingga kami dapat berpartisipasi di dalam Program Kreativitas Mahasiswa pada bidang Artikel Ilmiah (PKM-AI).

### **Kontribusi Penulis**

Penulis Satu yaitu Elvina Benedicta melakukan pengecekan plagiasi, melakukan analisis dan pemodelan, serta menyusun kesimpulan dan lampiran; Penulis Dua yaitu Wilbert Chandra melakukan performa algoritma, pemodelan, dan menyusun metode, hasil dan pembahasan, serta kesimpulan; Penulis Tiga yaitu Felix Cristiano Bungaran membuat visualisasi, serta menyusun ucapan terima kasih dan kontribusi penulis; Penulis Empat yaitu Patricia Pepita menyusun abstrak, pendahuluan, dan daftar pustaka serta melakukan pengecekan progres pembuatan artikel ilmiah; Peran dari Penulis Terakhir yaitu Ni Luh Putu Satyaning Pradnya Paramita, S.Si., M.Si., Ph.D adalah memberikan arahan serta saran dalam penyusunan artikel ilmiah.

### **Daftar Pustaka**

- Aprilliofany, A. 2020. Analisis penyebab pengangguran lulusan sekolah kejuruan di jawa barat dan garut. *Jurnal Litbang Sukowati*. 3 (2):12.
- Bachtiar, A.Z. 2019. Analisis pengaruh tenaga kerja informal, pengangguran, dan pendidikan terhadap pertumbuhan ekonomi di Indonesia. *Jurnal Ilmiah*. 5-6.
- Chakraborty, T., Chakraborty, A.K., Biswas, M., Banerjee, S. dan Bhattacharya, S. 2020. Unemployment rate forecasting: A hybrid approach. *Computational Economics*. 57:183-201.

- Hussain, M.S. 2023. *7 types of statistical distributions with practical examples*. URL: <https://datasciencedojo.com/blog/types-of-statistical-distributions-in-ml/>. Diakses tanggal 17 Februari 2024.
- Populix. 2023. *Statistik bayesian: Definisi, konsep utama, kegunaan*. URL: <https://info.populix.co/articles/statistik-bayesian/>. Diakses tanggal 17 Februari 2024.
- Prakoso, R.A. 2020. Analisis faktor-faktor yang mempengaruhi pengangguran terdidik di pulau jawa tahun 2010-2018. *Jurnal Ilmiah Mahasiswa FEB*. 9 (1):10.
- Pristiwanti, D., Badariah, B., Hidayat, S. dan Dewi, R.S. 2022. Pengertian pendidikan. *Jurnal Pendidikan dan Konseling*. 4 (6):2.
- Shabbir, A., Kousar, S. dan Zubair Alam, M. 2021. Factors affecting level of unemployment in South Asia. *Journal of Economic and Administrative Sciences*. 37 (1):1-25.
- Zangre, A. 2023. *Discrete vs. Continuous data: What's the difference?*. URL: <https://www.g2.com/articles/discrete-vs-continuous-data>. Diakses tanggal 17 Februari 2024.
- Zvornicanin, E. 2023. *Introduction to Gibbs Sampling*. URL: <https://www.baeldung.com/cs/gibbs-sampling>. Diakses tanggal 17 Februari 2024.

### Lampiran 1. Biodata Ketua dan Anggota, serta Dosen Pendamping

#### Biodata Ketua

##### A. Identitas Diri

1	Nama Lengkap	Patricia Pepita
2	Jenis Kelamin	Perempuan
3	Program Studi	Data Science
4	NIM	2602174176
5	Tempat dan Tanggal Lahir	Jakarta, 9 Mei 2004
6	Alamat E-mail	patricia.pepita@binus.ac.id
7	Nomor Telepon/HP	081293861222

##### B. Kegiatan Kemahasiswaan Yang Sedang/Pernah Diikuti

No	Jenis Kegiatan	Status dalam Kegiatan	Waktu dan Tempat
1	Google Developer Student Clubs (GDSC) BINUS Malang	Divisi Product & Curriculum: Data Science & Product Management	Oktober 2023 - Sekarang (Binus University)
2	Atma Jaya Consulting Community (ACC)	Vice President of Commercial	Juli 2023 - Sekarang (Jabodetabek)
3	Pengabdian Kepada Masyarakat (PKM) Data Science Club (DSC) BINUS University: DSC Goes to School	Pengajar	November 2023 (Binus University)
4	Student Consulting Project - Atma Jaya Consulting Community (ACC)	Student Consultant	Maret 2023 - Juni 2023 (Jabodetabek)
5	Asian English Olympic (AEO) BINUS English Club	Liaison Officer Divisi Transportation, Accommodation, and Food	Januari 2023 - Februari 2023 (Binus University)

##### C. Penghargaan Yang Pernah Diterima

No.	Jenis Penghargaan	Pihak Pemberi Penghargaan	Tahun
1	Best Consulting Team	Atma Jaya Consulting Community	2023
2	Juara 1 Kompetisi Menulis Esai	SMAK PENABUR Harapan Indah	2021
3	Juara 1 Karya Ilmiah Remaja	Schoters National Student Competition 2021	2021

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi.

Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan PKM-AI.

Jakarta, 17 - 02 - 2024

Ketua Tim



(Patricia Pepita)

**Biodata Anggota 1****A. Identitas Diri**

1	Nama Lengkap	Elvina Benedicta Santoso
2	Jenis Kelamin	Perempuan
3	Program Studi	Data Science
4	NIM	2602112322
5	Tempat dan Tanggal Lahir	Kediri, 26 Agustus 2004
6	Alamat E-mail	elvina.santoso@binus.ac.id
7	Nomor Telepon/HP	082228300576

**B. Kegiatan Kemahasiswaan Yang Sedang/Pernah Diikuti**

No	Jenis Kegiatan	Status dalam Kegiatan	Waktu dan Tempat
1	HIMTI	Aktivis	Februari 2023 - sekarang (Binus University)
2	Data Science Club	Member	April 2023 - sekarang (Binus University)
3			

**C. Penghargaan Yang Pernah Diterima**

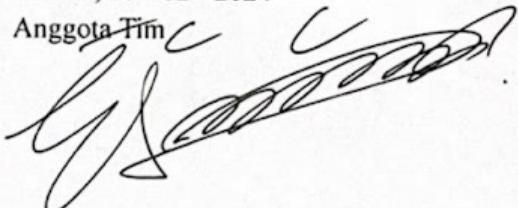
No.	Jenis Penghargaan	Pihak Pemberi Penghargaan	Tahun
1	Juara 1 Bridge Competition	Universitas Kristen Petra	2022
2	Juara 1 Interior Design Competition (INDESCO)	Universitas Ciputra Surabaya	2021 dan 2022
3	Juara Harapan 2 National Physics Competition (NuPho)	Universitas Negeri Malang	2021

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi.

Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan PKM-AI.

Jakarta, 17 - 02 - 2024

Anggota Tim



(Elvina Benedicta Santoso)

## Biodata Anggota 2

### A. Identitas Diri

1	Nama Lengkap	Wilbert Chandra
2	Jenis Kelamin	Laki-laki
3	Program Studi	Data Science
4	NIM	2602113666
5	Tempat dan Tanggal Lahir	Jakarta, 2 November 2004
6	Alamat E-mail	wilbert.chandra001@binus.ac.id
7	Nomor Telepon/HP	085155156773

### B. Kegiatan Kemahasiswaan Yang Sedang/Pernah Diikuti

No	Jenis Kegiatan	Status dalam Kegiatan	Waktu dan Tempat
1	Freshman Partner (FP)	Freshman Partner	Agustus 2023 - Sekarang (Binus University)
2	Data Science Club (DSC)	Member	Oktober 2023 - Sekarang (Binus University)
3	Bina Nusantara Computer Club (BNCC)	Member	September 2022 - Sekarang (Binus University)
4	Data Science Club (DSC)	Activist	April 2023 - Oktober 2023 (Binus University)
5	Bina Nusantara Computer Club (BNCC)	Activist	Oktober 2022 - Oktober 2023 (Binus University)
6	Technology Project Member (TPM) Bina Nusantara Computer Club	Back end member	Desember 2022 - Januari 2022 (Binus University)
7	Gojek Campus Ambassador	Ambassador	Oktober 2023 - Desember 2023 (Jabodetabek)

### C. Penghargaan Yang Pernah Diterima

No.	Jenis Penghargaan	Pihak Pemberi Penghargaan	Tahun
1			
2			
3			

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi.

Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan PKM-AI.

Jakarta, 17 - 02 - 2024

Anggota Tim



(Wilbert Chandra)

### Biodata Anggota 3

#### A. Identitas Diri

1	Nama Lengkap	Felix Cristiano Bungaran
2	Jenis Kelamin	Laki-laki
3	Program Studi	Data Science
4	NIM	2602170612
5	Tempat dan Tanggal Lahir	Jakarta, 15 Oktober 2004
6	Alamat E-mail	felix.bungaran@binus.ac.id
7	Nomor Telepon/HP	085811356531

#### A. Kegiatan Kemahasiswaan Yang Sedang/Pernah Diikuti

No	Jenis Kegiatan	Status dalam Kegiatan	Waktu dan Tempat
1	Himpunan Mahasiswa Teknik Informatika (HIMTI)	Activist	Februari 2023 - Sekarang (Binus University)
2	ISACA Student Group	Activist	Maret 2023 - Sekarang (Binus University)
3	Data Science Club	Member	April 2023 - Sekarang (Binus University)
4	Himpunan Mahasiswa Teknik Informatika (HIMTI)	Secretary of Kemanggisan	Februari 2024 - Sekarang (Binus University)

#### C. Penghargaan Yang Pernah Diterima

No.	Jenis Penghargaan	Pihak Pemberi Penghargaan	Tahun
1			
2			
3			

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi.

Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan PKM-KC.

Jakarta, 17 - 02 - 2024

Anggota Tim

(Felix Cristiano Bungaran)

### Biodata Dosen Pendamping

#### A. Identitas Diri

1	Nama Lengkap (dengan gelar)	Ni Luh Putu Satyaning Pradnya Paramita, S.Si, M.Si, PhD
2	Jenis Kelamin	Perempuan
3	Program Studi	Teknik Informatika
4	NIP/NIDN	0309059105
5	Tempat dan Tanggal Lahir	Madiun, 9 Mei 1991
6	Alamat E-mail	<a href="mailto:ni.paramita@binus.edu/">ni.paramita@binus.edu/</a> <a href="mailto:ni.luh@binus.ac.id">ni.luh@binus.ac.id</a>
7	Nomor Telepon/HP	081111129591

#### B. Riwayat Pendidikan

No	Jenjang	Bidang Ilmu	Institusi	Tahun Lulus
1	Sarjana (S1)	Statistika	Institut Teknologi Sepuluh Nopember	2012
2	Magister (S2)	Statistika	Institut Teknologi Sepuluh Nopember	2013
		Applied Mathematics	Aix-Marseille University	2013
3	Doktor (S3)	Operations Research	North Carolina State University	2023

#### C. Rekam Jejak Tri Dharma PT

##### Pendidikan/Pengajaran

No	Nama Mata Kuliah	Wajib/Pilihan	skls
1	Bayesian Data Analysis	Wajib	2
2	Introduction to Statistical Programming	Wajib	2
3	Survey and Sampling Methods	Wajib	2
4	Data Mining	Wajib	2

##### Penelitian

No	Judul Penelitian	Penyandang Dana	Tahun
1	Optimized polygenic risk scores for colorectal cancer		2024 - sekarang
2	Early detection of mental health issues in university students with machine learning	Binus University	2024 - sekarang
3	Interpretable machine learning for early detection	PT Meta Lab Nusantara	2023 - sekarang

	bipolar disorder based on crowdsourced symptoms		
4	Impact of law enforcement drug seizures on violent crimes and overdose mortality	National Science Foundation, USA	2022 - sekarang
5	Quantifying association and disparities between diabetes complications and COVID- 19 outcomes: A retrospective study using electronic health records	National Science Foundation, USA	2021 - 2023
6	Prediction of sepsis related mortality: an optimization approach	National Science Foundation, USA	2020 - 2021
7	Estimating city-level poverty rate based on e-commerce data with machine learning		2019-2020

**Pengabdian Kepada Masyarakat**

No	Judul Pengabdian kepada Masyarakat	Penyandang Dana	Tahun
1			
2			

Semua data yang saya isikan dan tercantum dalam biodata ini adalah benar dan dapat dipertanggungjawabkan secara hukum. Apabila di kemudian hari ternyata dijumpai ketidaksesuaian dengan kenyataan, saya sanggup menerima sanksi.

Demikian biodata ini saya buat dengan sebenarnya untuk memenuhi salah satu persyaratan dalam pengajuan PKM-AI.

Jakarta, 17 - 02 - 2024

Dosen Pendamping



(Ni Luh Putu Satyaning Pradnya  
Paramita, S.Si., M.Si., Ph.D)

**Lampiran 2. Kontribusi Ketua, Anggota, dan Dosen Pendamping**

No.	Nama	Posisi Penulis	Bidang Ilmu	Kontribusi
1.	Elvina Benedicta Santoso	Penulis pertama	Data Science	Melakukan pengecekan plagiasi, menyusun kesimpulan dan lampiran.
2.	Wilbert Chandra	Penulis kedua	Data Science	Melakukan performa algoritma dan menyusun metode, hasil dan pembahasan, serta kesimpulan.
3.	Felix Cristiano Bungaran	Penulis ketiga	Data Science	Membuat visualisasi, menyusun ucapan terima kasih dan kontribusi penulis.
4.	Patricia Pepita	Penulis keempat	Data Science	Menyusun abstrak, pendahuluan, dan daftar pustaka serta melakukan pengecekan progres pembuatan artikel ilmiah.
5.	Ni Luh Putu Satyaning Pradnya Paramita, S.Si., M.Si., Ph.D	Penulis terakhir	Statistika	Memberikan arahan dalam penyusunan artikel ilmiah.

### Lampiran 3. Surat Pernyataan Ketua Pengusul

#### **SURAT PERNYATAAN KETUA TIM PENGUSUL**

Yang bertanda tangan di bawah ini :

Nama Ketua Tim	:	Patricia Pepita
Nomor Induk Mahasiswa	:	2602174176
Program Studi	:	Data Science
Nama Dosen Pendamping	:	Ni Luh Putu Satyaning Pradnya Paramita, S.Si., M.Si., Ph.D
Perguruan Tinggi	:	Universitas Bina Nusantara

Dengan ini menyatakan bahwa proposal PKM-AI saya dengan judul: "PEMODELAN BAYESIAN UNTUK ANALISIS DAMPAK PENDIDIKAN TERHADAP TINGKAT PENGANGGURAN GLOBAL" yang diusulkan untuk tahun anggaran 2024 adalah:

1. Asli karya kami, belum pernah dibiayai oleh lembaga atau sumber dana lain
2. Tidak dibuat dengan menggunakan kecerdasan buatan/artificial intelligence (AI).

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan yang berlaku dan mengembalikan seluruh biaya yang sudah diterima ke kas Negara.

Demikian pernyataan ini dibuat dengan sesungguhnya dan sebenar – benarnya.

Jakarta, 17 – 02 – 2024

Yang menyatakan,



(Patricia Pepita)

2602174176

#### Lampiran 4. Surat Pernyataan Sumber Tulisan

##### **SURAT PERNYATAAN SUMBER TULISAN PKM-AI**

Saya yang menandatangani Surat Pernyataan ini:

Nama Ketua Tim	:	Patricia Pepita
Nomor Induk Mahasiswa	:	2602174176
Program Studi	:	Data Science
Nama Dosen Pendamping	:	Ni Luh Putu Satyaning Pradnya Paramita, S.Si., M.Si., Ph.D
Perguruan Tinggi	:	Universitas Bina Nusantara

1. Menyatakan bahwa PKM-AI yang saya tuliskan bersama anggota tim lainnya benar bersumber dari kegiatan yang telah dilakukan:
  - a. Sumber tulisan dari hasil kegiatan yang telah dilakukan berkelompok oleh tim penulis, yaitu: tugas kelompok dengan studi kasus
  - b. Topik Kegiatan: Analisis Model Linier Bayesian
  - c. Tahun dan Tempat Pelaksanaan: 2023, Jakarta Barat
2. Naskah ini belum pernah diterbitkan/dipublikasikan dalam bentuk prosiding maupun jurnal sebelumnya dan diikutkan dalam kompetisi.
3. Kami menyatakan kesediaan artikel ilmiah ini ditampilkan pada laman simbelmawa.

Demikian Surat Pernyataan ini dibuat dengan penuh kesadaran tanpa paksaan pihak manapun juga untuk dapat digunakan sebagaimana mestinya.

Jakarta, 17 - 02 - 2024

Yang menyatakan,



(Patricia Pepita)  
2602174176

**Lampiran 5. Hasil Pengecekan Plagiasi dengan Indeks Similaritas Maksimum 25%**

PatriciaPepita\_UniversitasBinaNusantara\_PKM-AI

---

ORIGINALITY REPORT

---



---

Gambar 7. Hasil pengecekan plagiasi